

Two-dimensional arrays with maximal complexity

MIRA-CRISTIANA ANISIU

Tiberiu Popoviciu Institute of Numerical Analysis

Romanian Academy

P.O. Box 68, 400110 Cluj-Napoca, Romania

e-mail: mira@math.ubbcluj.ro

and

ANTAL IVÁNYI

Faculty of Informatics of Eötvös Loránd University

H-1117 Budapest, Hungary, Pázmány Peter sétány 1/C.

e-mail: tony@inf.elte.hu

(Received: July 12–15, 2006)

Abstract. We present natural bounds for the complexity function of two-dimensional arrays, and we study the shape of the maximal complexity function. Some problems concerning the existence of maximal arrays are discussed.

Mathematics Subject Classifications (2000). 68R15, 05B20

1 Introduction

Arrays with elements from given sets of symbols have various applications, e.g. in frequency allocation of multibeam satellites [8], in designing mask configuration for spectrometers [12] and in cryptography [25]. In [22] possible applications in picture coding and processing are suggested.

Complexity is one of the important characteristics of such arrays. As the measure of the complexity we use the subarray complexity introduced by S. L. Ma [22] which is a natural generalization of the subword complexity defined by M. Heinz [13]. There are other measures as d -complexity [14, 19] or pattern complexity [6], and results in more dimensions [6, 9, 10, 15, 16, 22] too. One- and two-dimensional enumerative results can be found in [17].

In this paper we consider maximal and perfect arrays, and give conditions for their existence.

2 Definitions

Let $q \geq 2$ be a positive integer and $X = \{0, 1, \dots, q-1\}$ an alphabet. Let $X^{M \times N}$ denote the set of q -ary $M \times N$ arrays ($M, N \geq 1$ positive integers), and $X^{**} = \cup_{M, N \geq 1} X^{M \times N}$ the set of finite q -ary two-dimensional arrays.

DEFINITION 1 Let m and n be positive integers with $1 \leq m \leq M$ and $1 \leq n \leq N$. A q -ary $m \times n$ array $\mathcal{B} = [b_{ij}]_{m \times n}$ is a *subarray* of the q -ary $M \times N$ array $\mathcal{A} = [a_{kl}]_{M \times N}$ if there exist indices r, s such that $r+m-1 \leq M$, $s+n-1 \leq N$ and $b_{ij} = a_{r+i-1, s+j-1}$, $1 \leq i \leq m$, $1 \leq j \leq n$. \square

According to this definition only nonempty arrays can be (m, n) -subarrays.

We remark that we are dealing with *aperiodic* q -ary $M \times N$ arrays (written on a planar surface, with all the subarrays situated completely within the borders of the array). Another point of view is to consider the given array wrapped round on itself (written on a torus), hence a *periodic* array. Existence results for periodic and aperiodic arrays which contain every rectangular subarray of given sizes precisely once are given by Paterson [26], respectively Mitchell [24].

Notions of complexity similar to those for words can be introduced for arrays.

DEFINITION 2 Let $\mathcal{A} \in X^{M \times N}$ be a q -ary array and m, n positive integers with $1 \leq m \leq M$ and $1 \leq n \leq N$. Let $D_{\mathcal{A}}(m, n)$ denote the set of different $m \times n$ subarrays of \mathcal{A} . The *subarray complexity function*, or, simply, the *complexity function* $C_{\mathcal{A}} : \{1, 2, \dots, M\} \times \{1, 2, \dots, N\} \rightarrow \mathbb{N}$ of \mathcal{A} is

$$C_{\mathcal{A}}(m, n) = |D_{\mathcal{A}}(m, n)|, \quad m = 1, 2, \dots, M, \quad n = 1, 2, \dots, N, \quad (1)$$

and the *total complexity function* $T_{\mathcal{A}}$ of \mathcal{A} is

$$T_{\mathcal{A}} = \sum_{m=1}^M \sum_{n=1}^N C_{\mathcal{A}}(m, n). \quad \square \quad (2)$$

The one-dimensional complexity and total complexity functions were introduced by M. Heinz [13] in 1977, and studied later by many authors (see e.g. recent papers [1, 2, 3, 5, 11, 18, 19, 20, 21, 27]).

EXAMPLE 3 Let $X = \{0, 1, 2, 3, 4, 5\}$ be an alphabet and

$$\mathcal{A}_1 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad \mathcal{A}_2 = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 2 \end{pmatrix}, \quad \mathcal{A}_3 = \begin{pmatrix} 0 & 1 & 2 \\ 3 & 4 & 5 \end{pmatrix}.$$

Then $T_{\mathcal{A}_1} = 6$, $T_{\mathcal{A}_2} = 15$, and $T_{\mathcal{A}_3} = 18$. \square

DEFINITION 4 The q -ary $M \times N$ array \mathcal{A} is (q, m, n) -*extremal* if

$$C_{\mathcal{A}}(m, n) = \max_{\mathcal{B} \in X^{M \times N}} C_{\mathcal{B}}(m, n). \quad \square \quad (3)$$

DEFINITION 5 The q -ary $M \times N$ array \mathcal{A} is (q, m, n) -*perfect* if it contains each of the q^{mn} possible $m \times n$ q -ary arrays as a subarray exactly once. \square

DEFINITION 6 The arrays consisting of identical letters are called *homogeneous* arrays, the arrays consisting of different letters are called *rainbow* arrays. \square

We mention that a q -ary $M \times N$ rainbow array exists if and only if $q \geq MN$. It is obvious that (q, m, n) -extremal arrays always exist in $X^{M \times N}$ for arbitrary values of M, N , while (q, m, n) -perfect arrays can exist only for M, N satisfying $q^{mn} = (M - m + 1)(N - n + 1)$.

DEFINITION 7 The function $H_{q,M,N} : \{1, 2, \dots, M\} \times \{1, 2, \dots, N\} \rightarrow \mathbb{N}$ given by

$$H_{q,M,N}(m, n) = \min \{q^{mn}, (M - m + 1)(N - n + 1)\} \quad (4)$$

is called *maximal complexity function*. \square

DEFINITION 8 The q -ary $M \times N$ array \mathcal{A} is (q, m, n) -maximal if

$$C_{\mathcal{A}}(m, n) = H_{q,M,N}(m, n); \quad (5)$$

it is *maximal* if (5) holds for all $m = 1, 2, \dots, M$, $n = 1, 2, \dots, N$. \square

3 Bounds

We present the natural bounds of the complexity function for q -ary arrays $\mathcal{A} \in X^{M \times N}$, as well as those of the total complexity function.

PROPOSITION 9 For each q -ary $M \times N$ array \mathcal{A} we have

$$1 \leq C_{\mathcal{A}}(m, n) \leq \min \{q^{mn}, (M - m + 1)(N - n + 1)\}, \quad (6)$$

$$m = 1, 2, \dots, M, \quad n = 1, 2, \dots, N.$$

The lower bound is sharp for homogeneous arrays and the upper bound is sharp for rainbow arrays. The total complexity of \mathcal{A} satisfies the inequality

$$MN \leq T_{\mathcal{A}} \leq \sum_{i=1}^M \sum_{j=1}^N H_{q,M,N}(i, j). \quad (7)$$

Proof. From the definition of the subarray it follows that $C_{\mathcal{A}}(m, n) \geq 1$, $m = 1, 2, \dots, M$, $n = 1, 2, \dots, N$; for a homogeneous array the equality holds.

It is obvious that the complexity $C_{\mathcal{A}}(m, n)$ cannot exceed the total number of subarrays over X , that is q^{mn} ; it also cannot exceed the total number of subarrays of dimension $m \times n$ of the given array (possibly not all different), namely $(M - m + 1)(N - n + 1)$. It follows that $1 \leq C_{\mathcal{A}}(m, n) \leq \min \{q^{mn}, (M - m + 1)(N - n + 1)\}$, $m = 1, 2, \dots, M$, $n = 1, 2, \dots, N$. For a rainbow array \mathcal{R} we have $C_{\mathcal{R}}(m, n) = (M - m + 1)(N - n + 1) = \min \{q^{mn}, (M - m + 1)(N - n + 1)\}$.

By summing up the inequalities (6) we obtain (7). \blacksquare

REMARK 10 In terms of the maximal complexity functions, inequality (6) may be reformulated as

$$1 \leq C_{\mathcal{A}}(m, n) \leq H_{q,M,N}(m, n), \quad m = 1, 2, \dots, M, \quad n = 1, 2, \dots, N.$$

It follows that every (q, m, n) -perfect array, as well as any rainbow array, is (q, m, n) -maximal.

The values of the complexity and total complexity for homogeneous and rainbow arrays can be easily computed.

PROPOSITION 11 *If \mathcal{H} is a homogeneous $M \times N$ array and \mathcal{R} is an $M \times N$ rainbow array, then*

$$C_{\mathcal{H}}(m, n) = 1, \quad C_{\mathcal{R}}(m, n) = (M - m + 1)(N - n + 1),$$

$$m = 1, 2, \dots, M, \quad n = 1, 2, \dots, N$$

and

$$T_{\mathcal{H}} = MN, \quad T_{\mathcal{R}} = \frac{M(M+1)N(N+1)}{4}.$$

Proof. The complexity functions $C_{\mathcal{H}}$ and $C_{\mathcal{R}}$ were given in the proof of Proposition 9. Easy calculations give the formulas for $T_{\mathcal{H}}$ and $T_{\mathcal{R}}$. ■

The shape of the complexity function for words was proved in [21, 20, 18, 3] to be *trapezoidal*, i.e. it has an ascending part, possibly a horizontal one, and the last part is a descending line. The main feature is that after becoming constant, the complexity function of an arbitrary word cannot increase again. The question for arrays is: for a fixed m_0 , is $C_{\mathcal{A}}(m_0, \cdot)$ still trapezoidal? For $m_0 = 1$, the answer is positive, as a consequence of the mentioned result for words; nevertheless, this is not true for all the values $m_0 = 1, 2, \dots, M$. The array \mathcal{A} in the following example has the property that $C_{\mathcal{A}}(2, \cdot)$ increases again after becoming a constant.

EXAMPLE 12 *For the array $\mathcal{A} \in \{0, 1\}^{3 \times 19}$ given by*

$$\mathcal{A} = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 1 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 0 & 0 & 0 & 1 & 1 & 1 & 1 & 0 & 1 & 1 & 0 & 0 & 1 & 0 & 1 & 0 & 0 & 0 \end{pmatrix}$$

one has $C_{\mathcal{A}}(2, 4) = C_{\mathcal{A}}(2, 5) = 21$, $C_{\mathcal{A}}(2, 6) = C_{\mathcal{A}}(2, 7) = 22$ and $C_{\mathcal{A}}(2, 8) = 21$.

4 Properties of the maximal complexity function

We shall describe some properties of the function $H_{q,M,N}$ related to the shape of its graph, namely its monotonicity and its maximum.

For $M = 1$ (or $N = 1$) the arrays are in fact finite sequences (words). It was shown in [2, 3, 21] that for a given N we have

$$H_{q,1,N}(1, n) = \begin{cases} q^n, & n \leq k \\ N - n + 1, & k + 1 \leq n \leq N, \end{cases}$$

where k is the only natural number for which $q^k + k \leq N < q^{k+1} + k + 1$. The maximum of $H_{q,1,N}$ is equal to $N - k$ and is attained at the unique point $k + 1$ for $q^k + k < N \leq q^{k+1} + k + 1$, and at both k and $k + 1$ for $N = q^k + k$, hence $H_{q,1,N}$ is trapezoidal.

In the remaining part of this section we shall consider proper arrays (with $M, N \geq 2$).

REMARK 13 If both sizes of the array are smaller than the cardinal of the alphabet X ($M, N \leq q$), we have

$$(M - m + 1)(N - n + 1) \leq q^2 \leq q^{mn} \text{ for } mn \neq 1,$$

hence

$$H_{q,M,N}(m, n) = \begin{cases} \min\{q, MN\}, & m = n = 1 \\ (M - m + 1)(N - n + 1), & \text{otherwise.} \end{cases}$$

The maximum will be given by

$$H_{\max} = \max\{\min\{q, MN\}, N(M - 1), M(N - 1)\}$$

and will be attained at one of the points $(1, 1)$, $(1, 2)$ or $(2, 1)$. If $q < MN$, we have $H_{\max} = \max\{q, N(M - 1), M(N - 1)\}$; if $q \geq MN$, $H_{\max} = MN = h(1, 1)$.

In what follows we shall consider $\max\{M, N\} > q$.

PROPOSITION 14 *Let $m_0 \in \{1, \dots, M\}$ be fixed; the function $H_{q,M,N}(m_0, \cdot)$ is trapezoidal, the horizontal part containing at most two points; the last part is a descending line and the maximum of $H_{q,M,N}(m_0, \cdot)$ is attained at the first point d_{m_0} situated on the descending line, or on $d_{m_0} - 1$.*

Proof. The values of $H_{q,M,N}(m_0, n)$, $n \in \{1, \dots, N\}$ are given by the minimum of the values of an increasing exponential and of a descending line. At the beginning, if $(M - m_0 + 1)N > q^{m_0}$, $H_{q,M,N}(m_0, \cdot)$ will be situated on the exponential, and surely it will end on the descending line. Therefore $H_{q,M,N}(m_0, \cdot)$ will have a trapezoidal shape, with a horizontal part with at most two points.

There will be a point $d_{m_0} \leq N$ which is the least value of n for which $H_{q,M,N}(m_0, n)$ is on the descending line, i.e. if $d_{m_0} > 1$

$$\begin{aligned} (M - m_0 + 1)(N - d_{m_0} + 1) &\leq q^{m_0 d_{m_0}} \\ (M - m_0 + 1)(N - d_{m_0} + 2) &\geq q^{m_0(d_{m_0} - 1)}. \end{aligned}$$

The maximal value of $H_{q,M,N}(m_0, \cdot)$ will be given by

$$\mu_{m_0} = \max\left\{q^{m_0(d_{m_0} - 1)}, (M - m_0 + 1)(N - d_{m_0} + 1)\right\}.$$

The maximum of $H_{q,M,N}$ over $\{1, \dots, M\} \times \{1, \dots, N\}$ will be then $H_{\max} = \max\{\mu_m : m \in \{1, \dots, M\}\}$. ■

REMARK 15 The maximum of $H_{q,M,N}$ can be attained at a unique point (for example $H_{2,4,5}(2, 2) = 12$) or at several points ($H_{2,4,2}(1, 2) = H_{2,4,2}(2, 1) = H_{2,4,2}(3, 1) = 4$).

5 On the existence of maximal arrays

In [3] it was proved, using the results in [4, 7, 23, 27] that there exist finite words with maximal complexity, of any given length; it follows that there are $M \times 1$ and $1 \times N$ maximal arrays for all positive integers M and N . More than that, in [2] the number of the words with maximal complexity is presented. Nevertheless, if both M and N are ≥ 2 , the situation differs, as the following proposition shows.

PROPOSITION 16 *There are sizes $M, N \geq 2$ for which there are no arrays with maximal complexity.*

Proof. For $M = N = 4$ calculations show that the total complexity $T_{\mathcal{A}}$ of any 4×4 array is ≤ 69 , while $\sum_{i=1}^4 \sum_{j=1}^4 H_{2,4,4}(i, j) = 70$. It follows that for each 4×4 array \mathcal{A} there exists at least one pair (m, n) for which $C_{\mathcal{A}}(m, n) < H_{2,4,4}(m, n)$. ■

OPEN QUESTION Find the pairs M, N for which there exist maximal arrays in X^{**} .

The result in Proposition 16 prevents us from obtaining a q -ary array with maximal complexity for any M and $N \geq 2$. A further question is: given M, N and $m \leq M, n \leq N$, is there an $M \times N$ array $\mathcal{A}_{m,n}$ which is (q, m, n) -maximal?

A partial answer is given in [24]: in the binary case, if $(M-m+1)(N-n+1) = 2^{mn}$, there exists a $M \times N$ array which is $(2, m, n)$ -maximal (in fact it is $(2, m, n)$ -perfect).

Acknowledgment. The research of the first author was supported by the program GAR 13/2006 in the frame of academic exchange between the Hungarian Academy of Science and the Romanian Academy.

References

- [1] M.-C. ANISIU, *Finite words with special complexity properties*, Pure Math. Appl., **13** (2002), 31–37.
- [2] M.-C. ANISIU, Z. BLÁZSIK and Z. KÁSA, *Maximal complexity of finite words*, Pure Math. Appl., **13** (2002), 39–48.
- [3] M.-C. ANISIU and J. CASSAIGNE, *Properties of the complexity function for finite words*, Revue Anal. Numér. Théor. Approx., **33** (2) (2004), 123–139.
- [4] N.G. DE BRUIJN, *A combinatorial problem*, Proc. Kon. Ned. Akad. Wetensch., **49** (1946), 758–764.
- [5] A. CARPI and A. DE LUCA, *Words and special factors*, Theor. Comp. Sci., **259** (2001), 145–182.

-
- [6] J. CASSAIGNE, *Subword complexity and periodicity in two and more dimensions*. In: Developments in Language Theory, (Proc. of 4th Int. Conf. DLT, Aachen, 1999). Eds. G. Rosenberg and W. Thomas, World Scientific Publishing, River Edge, NJ, 2000, 14–21.
 - [7] L.J. CUMMINGS and D. WIEDEMANN, *Embedded de Bruijn sequences*, Congr. Numer., **53** (1986), 155–160.
 - [8] J. DÉNES and A.D. KEEDWELL, *Frequency allocation for a mobile radio telephone system*, IEEE Trans. Communications, **36** (1988), 765–767.
 - [9] T. ETZION, *Construction of perfect maps and pseudo-random arrays*, IEEE Trans. Inf. Theory, **34** (1988), 1308–1316.
 - [10] C.T. FAN, S.M. FAN, S.L. MA and M.K. SIU, *On de Bruijn arrays*, Ars Combin., **19A** (1985), 205–213.
 - [11] S. FERENCZI and Z. KÁSA, *Complexity of finite and infinite words*, Theor. Comp. Sci., **218** (1999), 177–195.
 - [12] M. HARVIT, *Spectrometric imager*, Appl. Optics, **10** (1971), 1415–1421.
 - [13] M. HEINZ, *Zur Teilwortkomplexität für Wörter und Folgen über einem endlichen Alphabet*, EIK, **13** (1977), 27–38.
 - [14] A. IVÁNYI, *On the d -complexity of words*, Annales Univ. Sci. Budapest., Sect. Comput., **8** (1987), 69–90.
 - [15] A. IVÁNYI, *Construction of infinite de Bruijn arrays*, Discr. Appl. Math., **22** (1988/89), 289–293.
 - [16] A. IVÁNYI, *Construction of three-dimensional perfect matrices*, Ars Combin., **29C** (1990), 33–40.
 - [17] A. IVÁNYI and Z. TÓTH, *Construction of planar de Bruijn words*. In: Papers on Automata and Languages, (ed. by I. Peák), pp. 63–69. Karl Marx Univ. of Economics, Budapest, 1988.
 - [18] S. JAEGER, R. LIMA and B. MOSSÉ, *Symbolic analysis of finite words*. In: The complexity function, Bull. Braz. Soc. Math. New Series **34** (2003), 457–477.
 - [19] Z. KÁSA, *On d -complexity of strings*, Pure Math. Appl., **9** (1–2) (1998), 119–128.
 - [20] F. LEVÉ and P. SÉÉBOLD, *Proof of a conjecture on word complexity*, Journées Montoises d’Informatique Théorique (Marne-la-Valée, 2000), Bull. Belg. Math. Soc. Simon Stevin, **8** (2001), 277–291.
 - [21] A. DE LUCA, *On the combinatorics of finite words*, Theor. Comp. Sci., **218** (1) (1999), 13–29.

- [22] S.L. MA, *A note on binary arrays with certain window property*, IEEE Trans. Inform. Theory, **30** (1984), 774–775.
- [23] M.H. MARTIN, *A problem of arrangements*, Bull. Amer. Math. Soc., **40** (1934), 859–864.
- [24] C.J. MITCHELL, *Aperiodic and semi-periodic perfect maps*, IEEE Trans. Inform. Theory, **41** (1995), 88–95.
- [25] C.J. MITCHELL and K.G. PATERSON, *Decoding perfect maps*, Des. Codes Cryptogr., **4** (1994), 11–30.
- [26] K.G. PATERSON, *Perfect maps*, IEEE Trans. Inform. Theory, **40** (1994), 743–753.
- [27] J. SHALLIT, *On the maximum number of distinct factors in a binary string*, Graphs Comb., **9** (2) (1993), 197–200.